



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

## Coronavirus

# Covid-19 still rife in Iran

An analysis of covid-19 infections among Iranian people casts further doubt on the idea that herd immunity can be achieved without vaccination, reports **Catherine Shaffer**

NEARLY everyone in Iran has been infected by the coronavirus at some point during the covid-19 pandemic, and some have caught the virus more than once, but the country still hasn't achieved herd immunity. Instead, Iran is seeing a punishing new wave of deaths driven by the delta variant.

Iran was one of the first countries after China to be hit by the pandemic, and it had a slow start to its vaccine roll-out. By July 2021, only about 3 per cent of the population was fully vaccinated against the coronavirus, according to Mahan Ghafari at the University of Oxford, although information from Johns Hopkins University in Maryland shows that proportion has now risen to about 23 per cent.

The Council on Foreign Relations, a US think tank, has called Iran's covid-19 mitigation measures "halting and ineffective", and said the country's response has been hampered by mixed messages from authorities.

Understanding the effects of covid-19 in Iran has been made more difficult by limited official data. Iran's Ministry of Health and Medical Education stopped releasing province-level data on confirmed cases and deaths in March 2020. Ghafari also says that low and middle-income countries are more likely to under-report cases and deaths due to a lack of health infrastructure and testing capacity.

To get round the shortfall in information, Ghafari and his colleagues analysed figures from Iran's National Organization for Civil Registration (NOCR) on how many deaths there were from all causes in the first nine months of 2020. They compared these with historical data to see how many more deaths there were than usual. The researchers used this excess deaths figure as a proxy to



**A student waiting to receive a covid-19 vaccine in Iran's capital Tehran**

**Firefighters disinfecting a street in Tehran**



estimate the number of covid-19 deaths and population-level exposure to the virus – an approach that has proved accurate in the UK and South Africa.

The NOCR began releasing past weekly data stratified by age group in August, which allowed the researchers to reconstruct the dynamics of the pandemic in Iran from January 2020 through to September 2021.

They calculated how many people in each province had contracted covid-19 using known global figures on the percentage of infected people who die, known as the infection fatality rate (IFR) for each age group. Dividing the number of excess deaths by the IFR gives an estimate for the number of infections, and allows you to tell what proportion of the population is infected, also known as the viral attack rate.

The analysis showed that the total infected population was probably very high in many provinces. Eleven of them had rates over 100 per cent as of 17 September. The highest rates were seen in Sistan and

Baluchestan province, which had an estimated attack rate of 259 per cent (medRxiv, doi.org/g2sn). If accurate, that would mean most people have had the virus twice and some for a third time. The researchers conclude that herd immunity through natural infection hasn't been attained in Iran in spite of widespread exposure to the SARS-CoV-2 virus, probably because of immunity waning over time, susceptibility to new variants of concern like delta or a combination of the two.

Mark Loeb at McMaster University in Canada says studies using direct measurements of infections are needed to conclusively determine the attack rate of the virus for a given population. However, he finds the study's conclusions plausible. "Overall, what's happening globally is that it certainly does appear that antibodies wane and there's not this point where, miraculously, herd immunity or herd effect is achieved," he says.

Loeb says he isn't aware of any pathogens for which herd immunity has been reached without vaccination. "By far the safest and most effective way to achieve herd immunity is through vaccination," he says.

The first strong indication that herd immunity against covid-19

**"There is no country so far, 19 months into the pandemic, that has reached herd immunity"**

can't be achieved without vaccination came from Manaus in Brazil. In October 2020, a study of blood donors there showed that 76 per cent of the population had been infected with the coronavirus, which was above the theoretical herd immunity threshold of 67 per cent for that

# Largest ever AI suggests limits to scaling up

Matthew Sparkes

community. That threshold is based on limiting the average number of other people an infected person is likely to go on to infect – the R number – to three. In December and January, Manaus saw an alarming surge in cases, hospitalisations and deaths, which was attributed to the highly transmissible gamma variant.

Evidence is piling up that the immunity conferred by infection wanes. A recent study of SARS-CoV-2 and related coronaviruses found that reinfection by SARS-CoV-2 occurred in about 16 months, less than half the time it takes for endemic “common cold” coronaviruses. And a study from India found a covid-19 reinfection rate of 27 per cent during the surge caused by the delta variant earlier this year (*Science*, doi.org/gm47sd).

David Fisman at the University of Toronto in Canada says, “Iran is a big and complicated place and I think the work [these researchers] have done has substantial face validity.” However, he says the study’s conclusions rest on two big assumptions: that all excess mortality is due to covid-19, and that the IFR estimates can be applied to Iran. He says if the IFR values aren’t the same there, because of overwhelmed health systems or a new, deadlier viral variant, the calculated attack rate would then also be inaccurate.

The analysis is consistent with previous studies showing that immunity to covid-19 fades and new variants aren’t stopped as effectively by immune defences, says Deepti Gurdasani at Queen Mary University of London. “There is no country so far, 19 months into the pandemic, that has reached herd immunity. And by that I mean, the point at which you lift all restrictions and the pandemic will actually die out because there’s nobody left to infect.” ■

MICROSOFT and chip maker Nvidia have created a vast artificial intelligence that can mimic human language more convincingly than ever before. But the cost and time involved in creating the neural network has called into question whether such AIs can continue to scale up.

The new neural network, known as the Megatron-Turing Natural Language Generation (MT-NLG), has 530 billion parameters, more than tripling the scale of OpenAI’s groundbreaking GPT-3 neural network that was considered the state of the art until now. This progress required more than a month of supercomputer access and 4480 high-power and expensive graphics cards, which are commonly used to run high-end neural networks.

When OpenAI released GPT-3 last year it surprised researchers with its ability to generate fluent streams of text. It had used 175 billion parameters – allocated slots of data within a computer that replicate the synapses between neurons in the human brain – and consumed vast amounts of text from which

to learn language patterns.

Microsoft and Nvidia, which released details of the AI in a blogpost, tested MT-NLG on a range of language tasks and found it could beat GPT-3, but not by much. On one benchmark, where an AI has to predict the end of a sentence, GPT-3 scored an accuracy of up to 86.4 per cent, while the new AI reached 87.2 per cent.

## 530 bn

Number of parameters that make up the new AI

This improved ability doesn’t come cheap. “It costs effectively millions of dollars to train one of these models,” says Bryan Catanzaro at Nvidia, as the computational resources needed to train it grow quickly.

MT-NLG was trained using Nvidia’s Selene supercomputer, which is made up of 560 powerful servers, each equipped with eight A100 80 gigabyte Tensor Core graphical processing units (GPUs). Each of those 4480 graphics cards – designed to run computer games, but also extremely capable at churning through vast amounts of data while training AIs – currently costs thousands of pounds

when bought commercially.

Even running the neural network once it is trained requires 40 of those GPUs, and each query takes between 1 and 2 seconds to process. This constant stretching of scale means that AI research is now, to a certain extent, an engineering problem of efficiently splitting up the problem and distributing it over vast amounts of hardware.

Catanzaro says that scale has been the dominant force in machine learning for decades. “It’s definitely true that better algorithms help, and it’s 100 per cent true that more data and better data absolutely helps, but I think that computing scale absolutely has been the driving force in a lot of progress,” he says.

Many researchers are reluctant to rely on scaling-up alone as they want a more elegant solution, says Catanzaro, but the results speak for themselves. Although the benchmark measurements reflect small improvements, there are thought to be significant steps up in the way the bigger AIs reason and extract nuanced information, which perhaps isn’t captured by ageing benchmarks, he says.

Samuel Bowman at New York University says that current benchmarks for assessing the quality of language processing AIs are nearing the end of their useful life. Researchers are seeking new metrics that can be used to assess the quality of language and even reasoning, he says, but that isn’t made simpler by the rapid rate of progress in AI.

Those same researchers are also “nervously waiting to find out” if scale can continue to bring improvements or whether it will hit a ceiling, he says, as the cost of research in the field grows rapidly. ■

Nvidia’s Selene supercomputer is a vast machine



NVIDIA